

Bull en Bear markten in verschillende Investeringsklassen



Abstract

Dit onderzoek is een uitbreiding van Hogerwerf et al (2009). Ik onderzoek of het mogelijk is aan de hand van een binair model te voorspellen wanneer investeringsklassen (aandelen, obligaties, grondstoffen en onroerend goed) bullish of bearish zullen zijn. Resultaat is dat het ongeveer 70 procent van de tijd mogelijk is de staat van een markt te bepalen. Een extra bivariate probit model wordt toegevoegd, waarin het mogelijk is de verschillende investeringsklassen te combineren, maar het blijkt dat toevoeging van dit model geen verbetering betekent.

Eindverslag Scriptie
Econometrie & Besliskunde
Jan Hogerwerf
31 juli 2009

Inleiding

Dit onderzoek schrijft voor op het werkcollege onderzoek 'Chasing Bulls' Fleeing Bears (Hogerwerf, van Opdurp, van Ovaa, Rooijen, 2009). In dit onderzoek is getracht een methode te ontwerpen om bull- en bearstaten voor de Amerikaanse S&P 500 index te voorspellen. Door meer te focussen op de grote lijn dan op individuele rendementen op korte termijn, hebben we getracht om meer stabiele returns te krijgen dan volgens een handelsstrategie met een korte horizon. We hebben dit gedaan door aan de hand van macro-economische en financiële indicatorvariabelen een logit model te schatten op bull en bearmarkten in de S&P 500. Hierna hebben we een handelsmodel opgezet, bestaande uit kopen van index futures wanneer het model een bull markt voorspelt en geld risicovrij wegzetten wanneer het model een bearmarkt voorspelt. Het lukt hierbij om de S&P 500 index met gemiddeld 5% te outperformen over een periode van 28 jaar.

Gezien de resultaten van dit onderzoek lijkt het interessant om dit onderzoek uit te breiden. Ik heb hierom besloten te kijken naar asset allocatie. Asset allocatie houdt in het verdelen van investering over verschillende investeringsklassen, regio's of landen. Voorbeelden hiervan zijn investeren in opkomende of traditionele landen, investeren in zwaar dividend betalende aandelen en groeiaandelen (value vs growth) en investeren in small en largecaps. Het idee hierachter is het vinden van de optimale risico / rendement afweging omdat dat de ene klasse de ene periode beter zal presteren dan de andere. Het idee van het verdelen van investeringen om zo de optimale portefeuille te vinden, komt voort uit de theorie van het CAPM (Sharpe 1964 en Black 1972), de best gediversifieerde portefeuille is degene waarin de volledige markt is opgenomen. Verschillende investeringsvormen zijn namelijk niet perfect gecorreleerd. Traditioneel wordt er bij asset allocatie voornamelijk gekeken naar de aandelenmarkt en de obligaties. Een voorbeeld hiervan is Li (2002). Deze kijkt naar correlatie tussen deze twee investeringsklassen en zoekt naar een verklaring door middel van macro-economische factoren. Deze focus op aandelen en obligaties is traditioneel zo omdat deze markten vrij liquide zijn en bovendien zorgen ze voor een periodieke uitbetaling in de vorm van dividend of rente. Door het combineren van deze klassen is het bijvoorbeeld mogelijk risico te reduceren. Voornamelijk vermogensbeheerders zoals pensioenfondsen gebruiken deze verschillende investeringsklassen om risico's weg te diversifiëren. Doordat markten door de jaren steeds meer liquide zijn geworden en transactiekosten sterk zijn gereduceerd, is de focus ook komen te liggen op andere markten dan deze traditionele twee investeringsklassen. Voorbeelden van deze klassen zijn grondstoffen, onroerend goed en buitenlands geld. Verschillend onderzoek is op dit gebied al gedaan, zo schrijven Bekkers, Doeswijk en Lam (2009) over correlaties tussen verschillende aandeelklassen. Hoevenaars et al (2007) kijken in hun onderzoek ook verder dan enkel de traditionele aandelen en obligaties. Zij ondervinden dat deze alternatieve investeringsklassen waarde kunnen toevoegen aan een portfolio en bovendien de eigenschap hebben risico's te diversifiëren.

In dit onderzoek zullen we echter niet alleen kijken naar asset allocatie om risico's te verkleinen. Door gebruik te maken van de indicatoren die we in het werkcollege

hebben gebruikt, zowel macro-economisch als financieel, ga ik bepalen of het mogelijk zal zijn een model op te zetten waarmee bull en bear markten in verschillende investeringsklassen correct kunnen worden voorspeld en of hiermee een succesvolle handelsstrategie valt op te zetten. Ik richt me hierbij op verschillende investeringsvormen, die ik evalueer door een representatieve benchmark te bekijken. Wellicht is het mogelijk om extra rendement te halen op investeringen door op het juiste moment in een markt te stappen. Men kan dan denken aan het veranderen van portefeuille samenstellingen per tijdseenheid, bijvoorbeeld per jaar of per maand. Dit is mogelijk doordat bijvoorbeeld omdat verschillende investeringsklassen verschillende business cycles hebben en dus bijvoorbeeld op elkaar voor of achter lopen. Deze conjunctuurcyclus maakt het onder andere mogelijk economische groei (deels) te voorspellen. Er zijn hier verschillende redenen voor (Dahlquist & Harvey, 2001), zo hebben consumenten de neiging inkomen snel te consumeren, investeringen bewegen echter minder snel, investeringsprojecten van bedrijven hebben vaak een looptijd van vijf jaar of langer.

In dit onderzoek concentreer ik me op aandelen, obligaties, onroerend goedsfondsen en grondstoffen. Om dit model op te zetten splits ik dit onderzoek op in verschillende delen. Allereerst is het nodig de juiste identificatie te vinden voor bull- en bearmarkten. Daarna moet er een voorspelmodel opgezet worden en tenslotte moeten de voorspellingen geevalueerd worden. Deze problemen komen uitgebreid aan bod in deze paper. De hoofdvraag is nu opgesplitst in de volgende deelvragen.

1. Hoe identificeer ik bull- en bearmarkten?
2. Welke indices gebruik ik en welke verklarende variabelen
3. Hoe bouw ik een voorspelmodel?
4. Hoe evalueer ik de voorspellingen?

In deze paper zal stapsgewijs getracht worden deze vragen te beantwoorden.

Data

In dit onderzoek kijk ik naar de volgende verschillende aandeleklassen: Aandelen (stocks), obligaties (bonds), grondstoffen (commodities) en onroerend goed (real estate). Voor aandelen gebruik ik de S&P 500 index, afkomstig van Robert Shiller, *Irrational Exuberance*¹. Voor obligaties neem ik de Barclays US Treasury index, afkomstig van datastream. Wat betreft commodities gebruik ik de S&P GSCI index, eveneens via datastream. Wat betreft vastgoed gebruik ik de NAREIT, North American composite price index, een index met daarin Amerikaanse beursgenoteerde vastgoedondernemingen, via de website van NAREIT². Deze data is volledig beschikbaar vanaf januari 1973, daarom begint de analyse op dit punt. Tabel 1 toont al wel aan dat deze beleggingsklassen niet bepaald perfect gecorreleerd zijn.

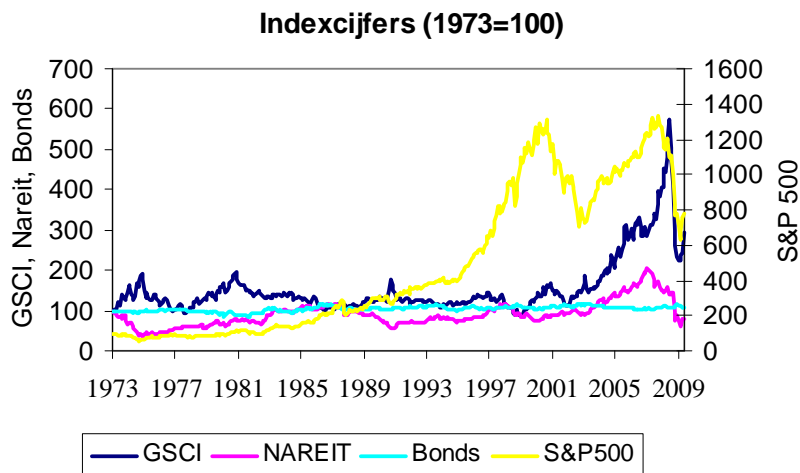
¹ <http://www.econ.yale.edu/~shiller/>

² www.reit.com

Tabel 1: correlaties van maandelijkse rendementen, 1973-2009

	Bonds	Gsci	S&P 500	Nareit
Bonds	1			
Csci	-0.063	1		
S&P 500	0.135	0.043	1	
Nareit	0.096	0.027	0.448	1

Figuur 1 toont de ontwikkeling van deze indices over tijd, waarin duidelijk wordt dat voornamelijk de S&P 500 een vlucht heeft genomen vanaf de jaren 80. De GSCI commodity index heeft ook een stijgende lijn het afgelopen decennia, gevolgd door een diep dal het in 2008. De andere twee indices fluctueren wel en hebben duidelijk opgaande periodes, gevolgd door neergaande, maar lijken toch behoorlijk stationair te zijn over de hele periode gezien



Figuur 1: prijsontwikkelingen indices, jan 1973=100

Wat betreft verklarende variabelen gebruik ik de data die gebruikt is in ons werkcollege verslag en in Chen (2008) die deze variabelen gebruikt om door middel van een univariaat regressie model voorspellingen uit te voeren, namelijk: Term spreads (het verschil tussen de 3-maands Treasury Bill Rate en de 10-jaars Treasury Constant Maturity Rate), inflatiepercentage, groei van de industriële productie, groei van de geldhoeveelheid (M1 en M2), werkloosheidspercentage, federal funds rate, verandering in gewogen wisselkoersindex, groei van de staatsschuld. Deze variabelen zijn afkomstig van de Federal Reserve of St. Louis. Bovendien gebruik ik prijs-earnings ratio van de S&P 500 (aandeelprijs gedeeld door winst per aandeel, Robert Shiller) en 1-month T-bill rate (Ibbotson and Associates, Kenneth R, French)

Tabel 2 toont deze variabelen schematisch en geeft testwaarden voor toetsen op stationariteit door middel van de Augmented Dickey-Fuller en Philips Perron test. Wat opvalt, is dat niet alle variabelen stationair zijn. Bij het gebruik van normale schattingsmethoden zou dit ongeldige resultaten opleveren. Volgens Park en Philips (2000) leidt het gebruik van niet-stationaire variabelen in een binaire regressie echter asymptotisch niet tot ongeldige resultaten, zolang er maar voldoende waarnemingen worden gebruikt dus.

Tabel 2: Unit root testen

Variabele	Transformatie	ADF		PP	
Federal funds rate	Niveau	-2.02		-1.82	
Industriële productie	% Verandering tov. Jaar geleden	-3.03	**	-3.58	**
Inflatie	% Verandering tov . jaar geleden	-1.52		-1.72	
M1	% Verandering tov . jaar geleden	-3.58	**	-2.64	*
M2	% Verandering tov . jaar geleden	-2.49		-2.79	*
Prijs/Earnings ratio	Niveau	-1.06		-1.11	
Staatsschuld	% Verandering tov . jaar geleden	-2.21		-2.06	
Term spread	Niveau	-3.47	**	-3.54	**
Werkloosheid	Niveau	-2.49		-2.07	
Wisselkoers	Verandering tov. maand geleden	-13.8	**	-13.6	**

Augmented Dickey-Fuller (ADF), Philips Perron (PP) zijn teststatistieken met als nulhypothese dat de serie een unit root heeft. Kritieke waarden zijn -2.87 (5%) -2.57 (10%).: Significant op 10% **: Significant op 5%..*

Door toegenomen efficiëntie in aandelenmarkten en het feit dat inefficiëntie niet per definitie leidt tot voorspelbare koersen, is het steeds lastiger geworden om met deze verklarende variabelen succesvol te voorspellen. Hogerwerf et al (2009) toonde echter aan, door deze variabelen op de juiste manier op te nemen en bovendien meer te focussen op de trend, dat er wel degelijk verklarende kracht zit in deze variabelen. Een kanttekening moet wel worden gemaakt bij de data. Het gebeurt regelmatig dat macro-economische data wordt aangepast, nadat deze beschikbaar zijn gekomen, bijvoorbeeld omdat de definitie van een cijfer wordt aangepast, of omdat er extra kennis beschikbaar is. Het zou hierdoor kunnen dat cijfers die ik gebruik anders zijn dan die op het relevante moment van voorspellen beschikbaar zijn. Helaas is het vaak lastig, zo niet onmogelijk om aan de eerste data te komen en andere studies nemen deze inconsistentie in data ook niet mee. Zodoende wordt hier geen correctie voor gemaakt, het zorgt wel voor een voorbehoud in de resultaten.

Methoden en Resultaten

Identificatie van bull- en bearmarkten

Wat betreft definitie van bull- en bearmarkt, gebruik ik dezelfde als in ons werkcollege, zoals die ook werd gebruikt in Lunde en Timmerman (2004), namelijk:

‘een aandelenmarkt switcht van een bullmarkt naar een bearmarkt en andersom als aandelenprijzen zijn gestegen/gedaald met een zeker percentage sinds hun vorige piek/dal in de huidige staat, oftewel een bepaalde drempelwaarde die wordt aangehouden.’

Het voordeel van deze definitie is dat hij ruimte laat voor negatieve prijsbewegingen tijdens een bullmarkt en positieve tijdens een bearmarkt, zolang de totale beweging vanaf de vorige piek maar niet groter is dan het drempelpercentage. Wat nodig is op te merken hierin is dat aandeelklassen zijn niet allemaal evenveel volatiliteit laten zien. Daarom kijk ik voor de verschillende klassen naar volatiliteit, door te kijken

wat voor volatiliteit de maandelijkse rendementen over tijd vertonen. Aan de hand hiervan kan ik bepalen voor welke reeks welk drempelpercentage correct is. Voor aandelen is het bijvoorbeeld vrij gebruikelijk om 20% aan te houden als drempelpercentage. Er zijn echter klassen veel minder volatiel en het heeft dus weinig zin om te wachten totdat deze gedaald zijn tot 20% van hun top (of dal). Aangezien de bonds een stuk minder volatiel zijn dan de andere klassen (Tabel 3), houd ik hiervoor een lager percentage aan van 10(10). Het 20(15) criteria (bear naar bull vs bull naar bear) bleek de beste voorspellingen af te geven in ons werkcollege, daarom concentreer ik me nu hier op.

	GSCI	Nareit	S&P 500	Bonds
μ	0.00424	0.000846	0.005765	0.000297
s	0.0595	0.0528	0.0454	0.0159

Tabel 3: gemiddelde en standaarddeviatie van maandelijkse returns op indices, periodieke uitbetalingen niet meegenomen

Uitvoering van dit identificatiealgoritme levert de volgende binaire reeksen op:

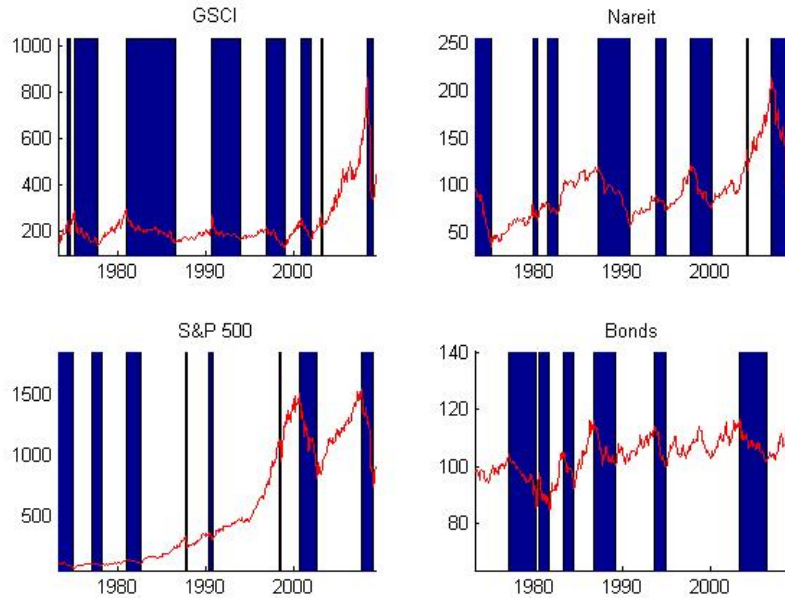
$$y_{j,t} = \begin{cases} 1 & \text{Als markt } j \text{ op tijdstip } t \text{ in bull staat} \\ 0 & \text{Als markt } j \text{ op tijdstip } t \text{ in bear staat} \end{cases}$$

Figuur 2 laat zien dat de de vier verschillende investeringsklassen niet op hetzelfde moment hun pieken en dalen vertonen. Wat in lijn ligt met de voorgaande bevindingen, is dat de S&P 500 zich voornamelijk in een bullmarkt bevindt vanaf begin jaren '80. Dit terwijl de andere vier investeringsklassen meer een cyclisch patroon vertonen, perioden van stijgende markten worden afgewisseld met perioden van dalende markten van ongeveer dezelfde lengte. Interessant is het wel om te zien dat, hoewel de Bond index stationair lijkt te zijn over tijd, deze index wel degelijk variatie over tijd vertoont die het in een bull en/of bearmarkt doet belanden.

Tabel 4 toont beschrijvende statistieken voor zowel bull als bearmaanden. Wat hierin opvalt, is dat in rendementen in bearmarkten over het algemeen meer pieken dan rendementen in bullmarkten. Wat we echter ook zien is dat bullstaten over het algemeen langer duren dan bearmarkten. Dit komt overeen met wat men over het algemeen ziet in rendementen, 'vertrouwen komt te voet en gaat te paard' en dat heeft zijn weerslag in de koersen

	GSCI		NAREIT		S&P500		Bonds	
	Bull	Bear	Bull	Bear	Bull	Bear	Bull	Bear
Gemiddeld rendement %	3.02	-2.25	1.65	-2.45	1.84	-2.73	0.34	-0.57
Max/min rendement %	26.19	-27.77	29.04	-30.82	16.30	-21.76	9.23	-5.76
Aantal bull perioden	16	15	11	11	11	11	7	6
Bull maanden	222	215	270	167	316	121	287	150
Gemiddelde duur	13.9	14.3	24.5	15.2	28.7	11.0	41.0	25.0

Tabel 4: statistieken voor zowel bull als bearmaanden



Figuur 2: Bull en Bearmarkten voor de vier investeringsklassen. De blauwe vlakken geven een bearmarkt aan

Het voorspelmodel

Er zijn twee conventionele manieren om voorspellingen uit te voeren op deze getransformeerde binaire reeksen, univariaat en multivariaat. Het verschil hierin is dat het in het multivariate model mogelijk is om correlaties tussen afhankelijke variabelen te modelleren, terwijl in het univariate model verondersteld wordt dat de correlatie tussen deze onafhankelijke variabelen 0 is. Het univariate model vormt dus een restrictie op het multivariate model. Een voor de hand liggende mogelijkheid voor het univariate model is een probit model te gebruiken voor de verschillende investeringsklassen. Op deze manier is het mogelijk de ongeconditioneerde kans bepalen dat een index zich in een bull of bearstaat gaat bevinden. Wat betreft het multivariate model is een mogelijkheid om een multivariaat probit model op te zetten die verschillende investeringsklassen kan combineren en daar mogelijk waardevolle informatie uit kan halen. Beide mogelijkheden worden in de volgende twee secties toegelicht.

Univariaat probit

Het voorspellen aan de hand van univariaat probit komt er op neer dat voor elke afzonderlijke investeringsklas een aparte voorspelling gemaakt wordt aan de hand van het identificatiecriterium dat we hebben besproken. Uitvoering van dit identificatie-algoritme levert voor elke index dus een binaire reeks op. Het is mogelijk hierop de volgende vergelijking te schatten:

$$P[y_{j,i} = 1] = F(x' \beta_i)$$

Waarbij

$$F(t) = \phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t \exp\left(-\frac{1}{2}s^2\right) ds$$

Een optimale oplossing kan worden gevonden door de log-likelihood van deze vergelijking te maximaliseren (zie Ben-Akiva en Lerman, 1985):

$$\text{Log}L(\beta|x) = \sum_{i=1}^n y_i \log(F(x'\beta_i)) + \sum_{i=1}^n (1 - y_i) \log(1 - F(x'\beta_i)) +$$

De integraal van de normale verdelingsfunctie is niet beschikbaar in gesloten vorm. Optimaliseren van dit probleem vereist zodoende numerieke optimalisatiesoftware. In de praktijk levert dit echter geen problemen op aangezien er tegenwoordig voldoende geavanceerde optimalisatiesoftware bestaat die een oplossing kunnen vinden voor deze vergelijking

Voor elke maand schat ik nu een afzonderlijk probit model, net zoals Pesaran en Timmerman (1995). In Hogerwerf et al (2009) werd voortdurend gebruik gemaakt van een moving window van 20 jaar (240 maanden). Omdat de beschikbaarheid van data een groter probleem is bij aandeelklassen anders dan de S&P 500, gebruik ik een moving window van 15 jaar (180 maanden). Dit is mogelijk zolang elke sample maar voldoende waarnemingen bevat van zowel bull als bear maanden. Wat hierbij moet worden opgemerkt, is dat het onmogelijk is om vanaf het begin van een beweging vanaf een lokale piek te weten of dit zal resulteren in een verandering van markt staat. Om dus een *look-ahead bias* te voorkomen, haal ik de maanden uit het model waarvan we nog niet weten of deze tot een bull- of bear markt behoren. Ik voorspel telkens één maand vooruit, beginnende in januari 1988 tot mei 2009. Omdat de verwachting is dat elke variabele in het model een invloed kan hebben, ga ik alle mogelijke combinaties van variabelen af. Dit betekent dat om te beslissen welk model de data het beste fit, er een selectiecriterium nodig is. In het werkcollege onderzoek bleken voornamelijk het Akaike Information Criterium (AIC) en het Schwarz Information Criterium (SIC) succesvolle modelselectiecriteria. Aangezien het verschil tussen deze twee gering is en de op basis van het SIC criterium gekozen modellen uit ons onderzoek de beste voorspelkracht liet zien, gebruik ik bij dit model alleen dit criterium voor modelselectie, met de volgende formule:

$$SIC(p) = \log(s_p^2) + \frac{p \log(n)}{n}$$

Waarin p het aantal regressoren is, s_p^2 de maximum likelihood schatter van de variatie in het model met p regressoren en n waarnemingen. In vergelijking met het AIC betekent dit dat het SIC over het algemeen modellen zal kiezen met iets minder variabelen.

Multivariaat probit

Het gebruik van multivariaat probit veronderstelt dat de correlatie tussen de verschillende klassen ongelijk aan 0 is. Het is op deze manier mogelijk de geconditio-

neerde verwachting uit te rekenen, de verwachte staat van een markt, gegeven de staat van een andere markt. Het probleem van multivariate verdelingsfuncties zit hem echter in n-variate normale verdelingsfunctie. Deze moet n-dimensionaal worden uitgerekend en dat kost veel tijd. In dit onderzoek combineer ik steeds twee klassen, wat het een bivariaat systeem maakt en waarin een oplossing wel uit te rekenen valt. De verschillende klassen kunnen dan op de volgende manier gecombineerd worden:

Wanneer we twee binaire variabelen hebben, Y_1 en Y_2 , waarvoor geldt:

$$\begin{cases} Y_1 = 1_{(y_1^* > 0)} \\ Y_2 = 1_{(y_2^* > 0)} \end{cases}$$

Dan is het model dat gespecificeerd wordt:

$$\begin{cases} Y_1^* = X\beta_1 + \varepsilon_1 \\ Y_2^* = X\beta_2 + \varepsilon_2 \end{cases} \quad \text{Met} \quad \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix} | X \approx N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right)$$

Uit deze modelvorm wordt onmiddellijk duidelijk dat twee univariate probit modellen een restrictie vormen op het algemene probit model, met restrictie $\rho=0$. Het is mogelijk de parameters van dit model te schatten aan de hand van Maximum Likelihood. Voor het afleiden van de likelihood functie volg ik Ashford en Sowden (1970): Wanneer we de samengestelde actie beschouwen van meer dan één systeem, dan is de kans dat zowel vergelijking y_1 als y_2 een zich in een bullmarkt bevinden:

$$P_{ij}^{11}(z) = F(x_i(z), x_j(z); \rho_{ij})$$

Waarbij $F(z)$ staat voor de bivariaat normale cumulatieve verdelingsfunctie met kansdichtheid:

$$f(y_1, y_2) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{y_1^2 - 2\rho xy_2^2}{2(1-\rho^2)}\right)$$

En ρ_{ij} gelijk is aan de correlatie tussen deze twee systemen, gecreëerd uit de residuen van twee enkelvoudige probit regressies op y_1 en y_2 .

De kans dat Y_1 zich in een bullmarkt bevindt, terwijl Y_2 dat niet is, is

$$P_{ij}^{10}(z) = P_i^1(z) - P_{ij}^{11}(z)$$

Met $P_i^1(z) = \Phi[x_i(z)]$, de cumulatief normale verdelingsfunctie. En op dezelfde manier:

$$P_{ij}^{01}(z) = P_i^1(z) - P_{ij}^{11}(z)$$

Voor de kans dat geen van beide klassen zich in een bullmarkt bevinden:

$$P_{ij}^{00}(z) = 1 - P_{ij}^{10}(z) - P_{ij}^{01}(z) - P_{ij}^{11}(z)$$

Op deze manier is het eenvoudig om de Log-likelihood bepalen,

$$L = \sum_j \sum_i r_{ij}^{st} \log(P_{ij}^{st}(z))$$

Met $r_{ij}^{st} = 1$ als de markten zich in de desbetreffende gecombineerde staat bevinden.

Optimalisatiesoftware maakt het mogelijk om deze likelihood functie te maximaliseren. Omdat het schatten van dit model veel tijd in beslag neemt, schat ik dit bivariate probit model steeds één keer met de vijf variabelen die het meest werden opgenomen in het univariate model en aan de hand van de geschatte parameters hiervan maak ik een voorspelling van $y_i = F(X'\beta)$, de kans dat de markt bullish zal zijn komende maand.

Evaluatie van voorspellingen

Het evalueren van de voorspellingen is op verschillende manieren mogelijk. Allereerst willen we graag kijken welke variabelen in het model opgenomen worden. Dit is interessant omdat er wellicht verschillende variabelen een voorspelkracht hebben op verschillende aandeelklassen. Daarna is het logisch om te kijken naar het percentage juiste voorspellingen. Dit doe ik door de voorspellingen in het enkelvoudig probit model, die bestaat uit een kans p_t te transformeren naar een binair getal door te zeggen dat $y_t=1$ als $p_t \geq c$ en $y_t=0$ als $p_t < c$. c is hierin het percentage bullmarkten in de voorspelsample, waarbij we dus elke keer een ander percentage gebruiken. Extra interessant is het hierbij om te onderzoeken of het bivariate probit model een verbetering is ten opzichte van het enkelvoudige probit model. Deze percentages kunnen worden vergeleken met het werkelijke aantal bullmaanden in de markt en het percentage maanden met een positief rendement. Ten slotte is het mogelijk om te kijken wat een strategie oplevert die belegt in deze investerings klas wanneer de markt bullish is en zich onthoudt van beleggen wanneer de markt bearish is.

Resultaten van het voorspelmodel

Het percentage opgenomen variabelen per investeringsklasse staan in Tabel 5 t/m 8. Gemiddeld worden er in het model rond de 6 van de 10 variabelen opgenomen. Opvallend zijn de diverse verschillen tussen de investeringsklassen. Zo zijn voornamelijk geldgroei, inflatie en werkloosheid zeer belangrijke voorspellers voor onroerend goed, voor grondstoffen is voornamelijk werkloosheid een belangrijke voorspeller. Voor bonds zijn federal funds rate en tevens geldgroei belangrijk, voor de S&P 500, geldgroei van M2, industriële productie en prijs/earnings ratio.

Percentages GSCI	Bear	Bull	Totaal
Inflatiepercentage	33.69	27.27	29.57
Federal funds rate	57.61	49.09	52.14
Groei Staatsschuld	68.48	58.79	62.26
Term spreads (3M-10Y)	56.52	41.82	47.08
Groei industriële productie	31.52	13.33	19.84
Groei M1	61.96	41.82	49.03
Groei M2	31.52	52.73	45.14
Verandering wisselkoersindex	3.26	3.03	3.11
Groei werkloosheid	94.57	98.18	96.89
Prijs/earnings ratio	30.35	35.4	33.60
Gemiddeld aantal variabelen	4.69	4.21	4.39

Tabel 5: opgenomen variabelen voor de GSCI op basis van het Schwarz Informatie Criterium

Percentages NAREIT	Bear	Bull	Totaal
Inflatiepercentage	87.69	82.67	85.21
Federal funds rate	87.69	43.31	65.75
Groei Staatsschuld	65.75	79.52	71.21
Term spreads (3M-10Y)	52.3	48.81	50.58
Groei industriële productie	29.23	70.86	49.8
Groei M1	83.07	92.91	87.93
Groei M2	25.38	31.49	28.4
Verandering wisselkoersindex	3.07	2.36	2.72
Groei werkloosheid	88.46	68.5	78.6
Prijs/earnings ratio	84.78	55.16	65.75
Gemiddeld aantal variabelen	6.07	5.76	5.86

Tabel 6: opgenomen variabelen voor de NAREIT op basis van het Schwarz Informatie Criterium

Percentage S&P500	Bear	Bull	Totaal
Inflatiepercentage	78.1	33.55	51.75
Federal funds rate	24.76	61.18	46.3
Groei Staatsschuld	54.28	61.18	58.36
Term spreads (3M-10Y)	36.19	50	44.35
Groei industriële productie	60	69.74	65.76
Groei M1	22.86	17.11	19.46
Groei M2	72.38	72.37	72.37
Verandering wisselkoersindex	12.38	25.66	20.23
Groei werkloosheid	38.1	40.79	22.18
Prijs/earnings ratio	54.28	71.71	62.64
Gemiddeld aantal variabelen	4.53	5.03	4.63

Tabel 7: opgenomen variabelen voor de Bonds op basis van het Schwarz Informatie Criterium

Percentages Bonds	Bear	Bull	Totaal
Inflatiepercentage	9.33	34.06	26.84
Federal funds rate	94.67	68.68	76.26
Groei Staatsschuld	65.33	57.69	59.92
Term spreads (3M-10Y)	59.92	34.62	36.96
Groei industriële productie	70.67	73.08	72.37
Groei M1	88.00	91.208	90.27
Groei M2	65.33	47.8	52.92
Verandering wisselkoersindex	0	0	0
Groei werkloosheid	70.67	46.15	53.31
Prijs/earnings ratio	73.33	55.49	60.7
Gemiddeld aantal variabelen	5.97	5.09	5.30

Tabel 8: opgenomen variabelen voor de Bonds op basis van het Schwarz Informatie Criterium

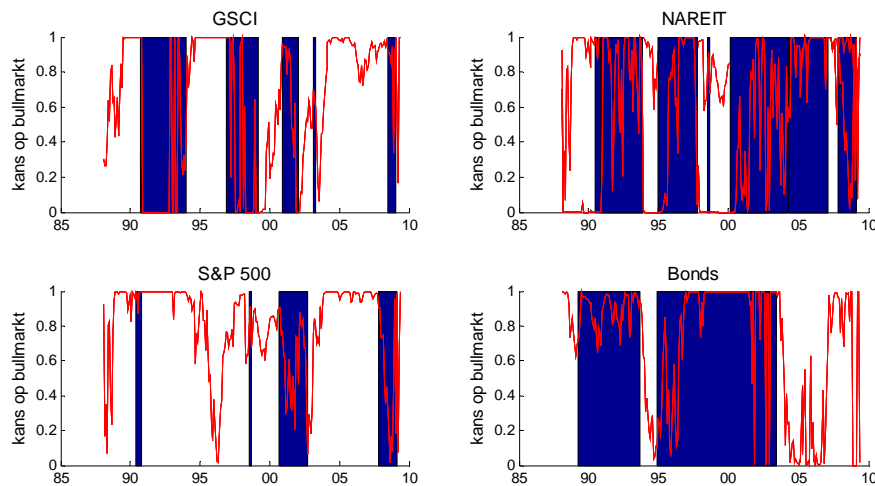
Wat betreft de voorspelkracht van het enkelvoudig probit model, staan de uitkomsten in Tabel 9. Om te kijken hoe het model presteert heb ik de regressies ook gedaan zonder verklarende variabelen, dus met enkel een constante.

	S&P 500	Nareit	GSCI	Bond
% maanden met een positief rendement	62.26	53.70	54.47	52.14
Werkelijk % bullmaanden	81.32	57.59	60.31	73.93
% correct voorspeld	66.92	70.81	72.31	78.98
% incorrect voorspeld	33.08	29.19	27.69	21.02
Voorspelling constante	40.46	49.02	57.19	28.40
Voorspelmaanden	257	257	257	257

Tabel 9: percentages correcte en incorrecte voorspellingen voor de vier verschillende investeringsklassen

Goed zichtbaar is het dat het voorspelmodel aan alle voorspelmodellen iets extras toevoegt. Alleen voor de S&P 500 blijven de voorspellingen iets achter bij het werkelijke percentage bullmaanden, maar lopen ze wel vooruit op een voorspelling met alleen een constante. Dit is ondermeer opvallend omdat uit Hogerwerf et al (2009)' wel degelijk bleek dat aan de hand van dit model een grote verbetering verwezenlijkt kon worden. Een verklaring hiervoor kan zijn het gebruik van een probit model in plaats van een logit model en bovendien de kleinere samples waar ik mee schat en voorspel. Figuur 3 toont de voorspelde kansen voor een bullmarkt voor de vier verschillende investeringsklassen. Zichtbaar is dat over het algemeen de marktstaat heel aardig voorspeld kan worden, enkel wanneer de markt zich langere tijd in een bepaalde staat bevindt dan vinden er onverwachte wisselingen plaats. Dit hoeft niet vreemd te zijn, het is goed mogelijk dat de markt wel verandert, maar dat het drempelpercentage net niet wordt gehaald, waardoor er in werkelijkheid net geen regime switch plaatsvindt.

Figuur 3: voorspelde kansen univariate model. Blauwe vlakken geven bearmarkten aan



	GSCI	Nareit	S&P 500	Bond
Saldo handelsstrategie	2.31	2.98	6.078	1.24
Saldo buy and hold strategie	6.789	-0.214	1.941	0.077

Tabel 10: rendementen aan de hand van een eenvoudige handelsstrategie, dividenden, periodieke uitbetalingen en transactiekosten niet meegenomen

Tabel 10 toont de rendementen van een eenvoudige handelsstrategie wanneer geïnvesteerd wordt in de desbetreffende investeringsklasse wanneer het model daar de aanleiding toe geeft en risicovrij het geld wordt weggezet wanneer het model een bearmarkt voorspelt. Wat opvalt is dat alle rendementen hoger zijn dan dat ze zouden zijn wanneer een buy and hold strategie wordt aangehouden. Doordat het rendement over de hele periode van de S&P 500 index echter zo veel hoger is dan de andere indices, is het lastig om hieruit een handelsstrategie op te maken, zeker omdat we dan ook dividenden en rentes mee zouden moeten nemen. Obligaties en Onroerend goed worden getypeerd doordat een groot deel van het rendement op deze investeringsklassen wordt behaald door deze uitbetalingen en dan zou hun rendement een stuk laten stijgen.

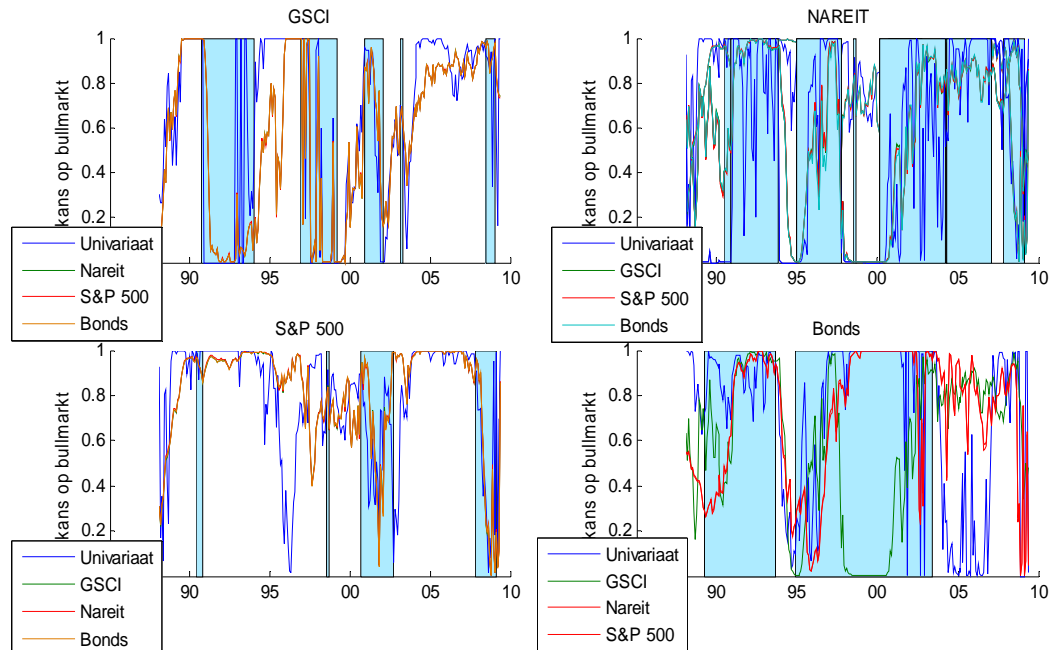
Resultaten bivariaat model

In Tabel 11 staan de uitkomsten van de voorspellingen aan de hand van het bivariaat probit model. Het blijkt dat de voorspellingen wel iets afwijken van de voorspellingen uit het univariate model. Dit kan ten eerste komen doordat het model minder en andere variabelen meeneemt en dat de winst van het bivariate model, het meenemen van een andere index, teniet wordt gedaan door dit verlies aan verklarende variabelen. Interessant om te zien is ook dat het combineren van de verschillende indices weinig uitmaakt voor de kwaliteit van de voorspellingen. Figuur 4 laat ook nog eens zien dat de voorspelde kansen van de univariate en bivariate probit modellen, zeker die van de verschillende bivariate modellen, niet echt veel uit elkaar liggen. Waarschijnlijk ligt de verklaring hierin dat de correlatie tussen de verschillende investeringsklassen niet echt groot is en dus is er weinig extra rendement te behalen door deze klassen te combineren. Bovendien zijn de beste voorspellers uit het univariate model opgenomen in het bivariate model, wat maakt dat de meeste informatie die in het univariate model zit, ook in het bivariate model aanwezig is.

Voorspelling,gegeven	GSCI	Nareit	S&P 500	Bonds
GSCI	-	71.21	71.21	70.82
Nareit	70.82	-	70.43	70.43
S&P 500	69.65	70.43	-	70.04
Bonds	63.42	64.59	64.59	-

Tabel 11: percentage juist voorspelde maanden aan de hand van het bivariaat probit model

Figuur 4: voorspelde kansen voor de verschillende indices. Lichtblauwe vlakken geven een bearmarkt aan



Conclusie

Doel van dit onderzoek was te onderzoeken of het mogelijk is het model dat gebruikt wordt in Hogerwerf et al (2009) om bull en bear markten voor de S&P 500 te voorspellen toe te passen op andere investeringsklassen, te weten aandelen, onroerend goed, obligaties en grondstoffen. Ik heb dit gedaan door aan de hand van de definitie over bull en bear markten die in het werkcollege is gebruikt, zowel een enkelvoudig probit model te schatten voor deze investeringsklassen apart als de investeringsklassen te combineren tot een bivariaat probit model.

Wat gebleken is uit dit onderzoek, is dat er verklarende kracht zit in dit model ook voor de verschillende investeringsklassen. Wat is opgevallend dat de variabelen die voorspellende kracht blijken te hebben nogal verschillen tussen de verschillende investeringsklassen. Omdat het rendement op de S&P500 in deze sample zo veel hoger is geweest dan op de andere investeringsklassen, is het echter lastig om het rendement te verhogen. Wanneer we hier een serieus handelsmodel voor zouden willen opzetten zouden we periodieke uitbetalingen ook in het model op moeten nemen. Het bivariate probit model voegt nauwelijks voorspellende kracht toe aan het model en gezien de grote hoeveelheid rekentijd en complexiteit die het met zich meebrengt, die ervoor zorgt dat we minder verklarende variabelen op kunnen nemen, is het gebruik ervan geen toevoeging aan het model dat we aanvankelijk hebben gebruikt. Nieuw onderzoek zou zich wellicht beter kunnen spitsen op het vinden van nieuwe indicatoren of om de investeringsklassen op een andere manier in het model op te nemen. Extra onderzoek zal nog moeten uitwijzen of met behulp van dit model een combinatiestrategie kan worden opgezet waarbij het mogelijk is om te investeren in de investeringsklasse die de hoogste kans heeft om bullish te worden of het hoog-

ste verwachte rendement heeft, maar dan zal wel moeten onderzocht of het mogelijk is om hierin het effect van transactiekosten en periodieke uitkeringen mee te nemen.

Referenties

Ashford J.R., R.R. Sowden, Multi-Variate Probit Analysis, *Biometrics* Vol. 26, No.3, p535-546. september 1970

Ben-Akiva M., M., S.R. Lerman, Discrete Choice analysis. Cambridge, MA: MIT Press, 1985

Black, F., C. Jensen, M. Scholes, The Capital Asset Pricing Model, some empirical tests, 1972

Bekkers, Doeswijk, Lam: Strategic Asset Allocation: Determining the optimal portfolio with ten asset classes, maart 2009

Chen, S.-S. (2009). Predicting the bear stock market: Macroeconomic variables as leading indicators. *Journal of Banking & Finance*, 33:211–223.

Dahlquist, M., C.M. Harvey, Global Tactical Asset Allocation, januari 2001

Hoevenaars, R., R. Molenaar, P. Schotman, T. Steenkamp, Strategic Asset Allocation with liabilities: beyond stocks and bonds. *Journal of Economic Dynamics & Control* 32 (2008) 2939–2970

Hogerwerf, J., J. van Opdurp, T. Ovaa, M. van Rooijen, Chasing Bulls, Fleeing Bears, *Werkcollege Bachelor Econometrie & Besliskunde*, Erasmus Universiteit Rotterdam, mei 2009

Li, L.: Macroeconomic Factors and the Correlation of Stock and Bond Returns, Yale ICF Working Paper No. 02-46; AFA 2004 San Diego Meetings, november 2002.

Lunde, A. & A. Timmermann (2004). Duration dependence in stock prices: An analysis of bull and bear markets. *Journal of Business & Economic Statistics*, 22(3):253–273.

Park & Philips (2000). Nonstationary binary choice. *Econometrica*, 68(5):1249–1280.

Sharpe, W.F., Capital Asset Prices: A theory of market equilibrium under conditions of risk, *The journal of finance*, Volume 19, no.3, september 1964